

# MiniJava Compiler Evaluation

Daniel Grund

`grund@cs.uni-saarland.de`

November 3, 2011

## Popular errors

- ▶ Extraneous output
- ▶ Forgotten keywords: e.g. `true`, `false`
- ▶ `//` is not a comment in MiniJava
- ▶ Wrong handling of comments
- ▶ Trouble with EOL / EOF
- ▶ Fixed size assumptions: lines/idents are shorter than 1024 chars

## Design / High-level

- ▶ Process the input character by character (not per line)
- ▶ You don't need to explicitly store the list of tokens
- ▶ Do not hard-code `printfs` into the lexer
- ▶ If you detect a bug, write a test case for it
- ▶ Write more test cases, really, I mean it
- ▶ If something is unclear, ask

## Programming / Testing

- ▶ Using `boost::regex` or similar is not allowed
- ▶ Take C/C++ warnings as errors
- ▶ Concerning EOF, read the documentation of `std::ifstream::get()` or `fgetc()`, respectively
- ▶ Uncaught throw-ups will *not* yield return code 1

ID  $\mapsto$  Group

2031946	P	2519866	P	2526193	H	2531557	R
2511042	C	2520925	O	2526240	K	2531835	A
2511081	P	2521139	A	2526241	K	2535203	V
2513331	B	2522025	D	2526415	H	2535483	
2513509	E	2522705	I	2526486	F	2536539	W
2514322	G	2522718	N	2527140	F	2536586	W
2514359	C	2523081	B	2528251	F	2539125	T
2515294	E	2523114	J	2528501	O	2539149	T
2516088	B	2523129	N	2528532	R	2539152	T
2517200	Q	2523343	N	2530270	R	2539755	W
2518099	I	2524010	L	2530469	M	2539760	U
2518162	Q	2524214	J	2530559	A	2539954	U
2518183	J	2524265	J	2530690	M	2540322	G
2518561	O	2524568	B	2530699	M	2540324	G
2518686	E	2524649	L	2530711	V	2540414	S
2518834	E	2525627	I	2530947	S	2540807	T
2518894	D	2525783	D	2531407	M	2540817	V
2519766	S	2526159	H	2531460	A		

# Correctness Tests: Lexer

test case	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	Z
all_chars.java	F			A	X	-			F			F		F	F				F	X	S	-	F	
all_tokens.java	F			A	X	-		X		X		F			X				F	X	S	-	F	
big_number.java				X	X	-														X	S	-	F	
Bottles.java				X	X	-														X	S	-	F	
char_00.java		E		E		-				255								S		E	S	-		
char_0D.java					X	-			F		F			F						X	S	-	F	
char_FF.java				E	E	-		E	E	255	E									E	S	-		
char_invalid.java				A		-				255										-	S			
comment_2_in_1_line.java					X	-														X	S	-	F	
comment_at_end.java					X	-														X	S	-	F	
comment_at_start.java					X	-														X	S	-	F	
comment_minimal.java					X	-														-	S			
comment_multi-line.java					X	-														X	S	-	F	
comment_only.java					X	-														-	S			
comment_terminated.java	I			E		-	A			255	E			E		E				E	S	-		
comment_with_inv_char.java					X	-														X	S	-	F	
comment_with_slash.java					X	-		F												X	S	-	F	
double_slash.java					X	-	X						X		X					X	S	-	F	
empty.java					X	-					F									-	S			
exercise_example.java					X	-														X	S	-	F	
Fac.java				X	X	-														X	S	-	F	
Fib.java				X	X	-														X	S	-	F	
GCD.java				X	X	-														X	S	-	F	
HelloSomething.java				X	X	-														X	S	-	F	
identifiers.java	F				X	-												X	F	X	S	-	F	
long_ident.java					X	-												X		X	S	-	F	
MinInt.java				X	X	-														X	S	-	F	
mixed_eols.java					X	-		F		F				F						X	S	-	F	
no_eol.java					X	-										X				X	S	-	F	
numbers.java					X	-														X	S	-	F	
token_pairs.java	F			A	X	-	X	X	F	F			F	X		X			F	X	S	-	F	
tokens_seperated_by_comment.java					X	-														X	S	-	F	
whitespace.java					X	-														X	S	-	F	
passed (of 33)	28	32	33	19	32	4	0	31	25	27	29	27	33	29	27	28	0	28	29	3	0	0	4	33
	10,282	13,456	12,593	27,491	7,840	7,557	0	16,491	13,043	6,557	8,967	9,737	23,417	16,675	4,017	9,354	0	103,252	34,827	12,958	4,662	0	5,210	46,121

size (bytes)

X wrong output  
 E error not detected  
 F failed with error on valid input  
 S segfault  
 A assertion/abort  
 I infinite loop

# Throughput Test: Lexer

